

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Song Z, Rose S, Safran DG, Landon BE, Day MP, Chernew ME. Changes in health care spending and quality 4 years into global payment. N Engl J Med 2014;371:1704-14. DOI: 10.1056/NEJMsa1404026

Supplementary Appendix

Changes in Health Care Spending and Quality 4 Years into Global Payment

Zirui Song, M.D., Ph.D., Sherri Rose, Ph.D., Dana Gelb Safran, Sc.D., Bruce E. Landon, M.D., M.B.A., Matthew P. Day, F.S.A., M.A.A.A., Michael E. Chernew, Ph.D.

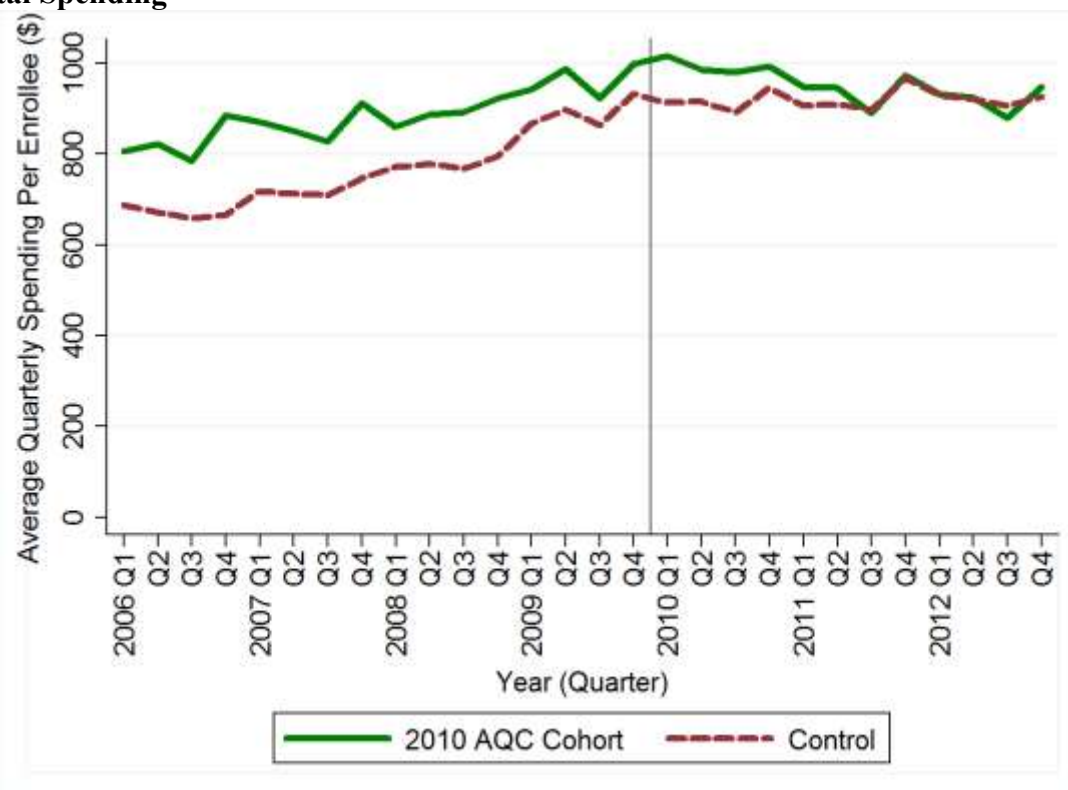
Table of Contents

Figure S1.	Unadjusted Spending: 2010 AQC Cohort vs. Control A. Total Spending B. Decomposition By Site and Type of Care
Figure S2.	Unadjusted Spending: 2011 AQC Cohort vs. Control A. Total Spending B. Decomposition By Site and Type of Care
Figure S3.	Unadjusted Spending: 2012 AQC Cohort vs. Control A. Total Spending B. Decomposition By Site and Type of Care
Figure S4.	Process Quality by AQC Cohort, Aggregate Results 2007-2012 A. Chronic Care Management B. Adult Preventive Care C. Pediatric Care
Table S1.	Ambulatory Quality Measures in the Alternative Quality Contract
Table S2.	Provider Organizations in the Alternative Quality Contract
Table S3.	Characteristics of the Population: AQC Cohorts vs. Control A. 2009 AQC Cohort vs. Control B. 2010 AQC Cohort vs. Control C. 2011 AQC Cohort vs. Control D. 2012 AQC Cohort vs. Control
Table S4.	Change in Average Quarterly Spending per Enrollee, AQC Cohorts vs. Control
Table S5.	Adjusted Spending: 2009 AQC Cohort Subgroups vs. Control*

Table S6.	Sensitivity Analyses
	A. Alterations to the Statistical Model
	B. Alterations in Variables or Sample
Table S7.	HEDIS New England Averages
	A. Process Measures
	B. Outcome Measures
Table S8.	Outcome Quality: AQC Cohorts vs. HEDIS National Average
	A. 2009 AQC Cohort vs. Control
	B. 2010 AQC Cohort vs. Control
	C. 2011 AQC Cohort vs. Control
	D. 2012 AQC Cohort vs. Control

Figure S1. Unadjusted Spending: 2010 AQC Cohort vs. Control.*

A. Total Spending



B. Decomposition By Site and Type of Care

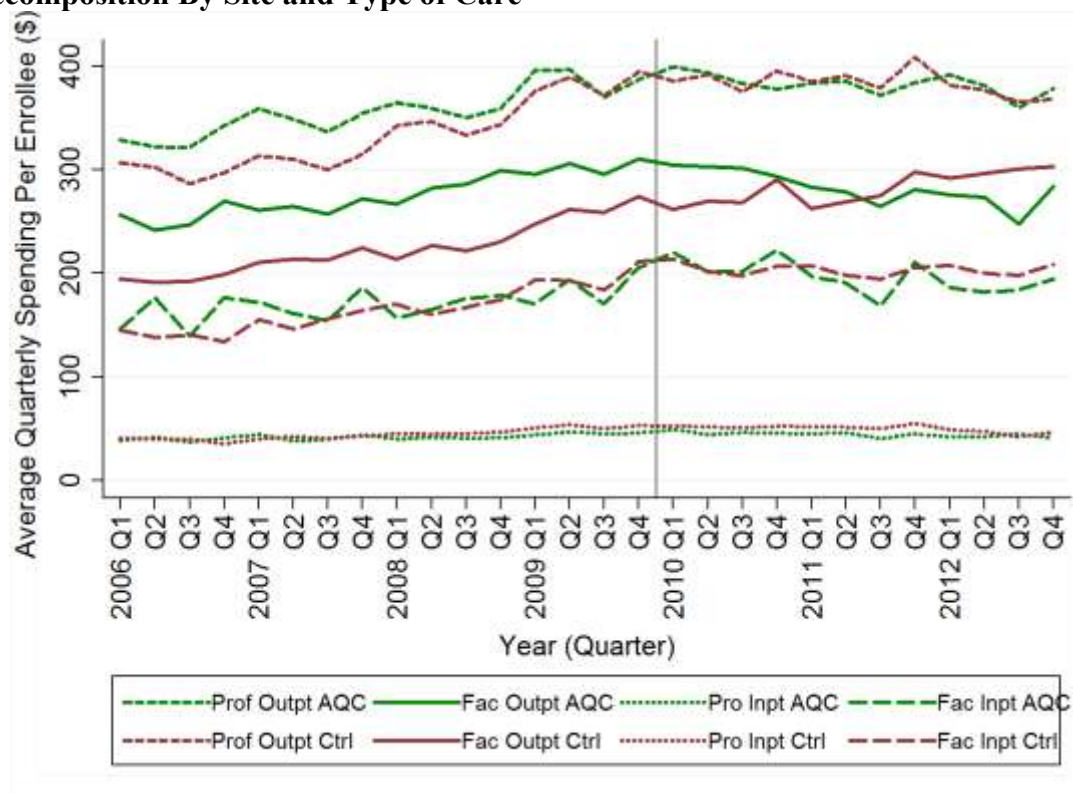
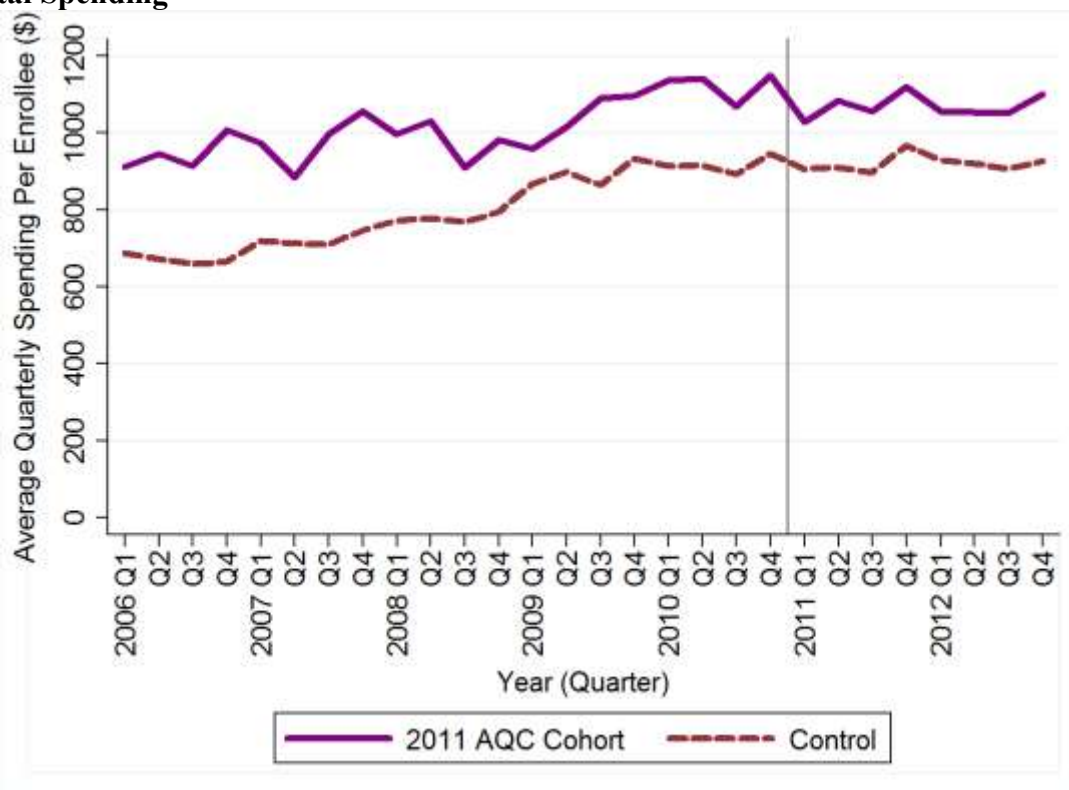


Figure S2. Unadjusted Spending: 2011 AQC Cohort vs. Control.*

A. Total Spending



B. Decomposition By Site and Type of Care

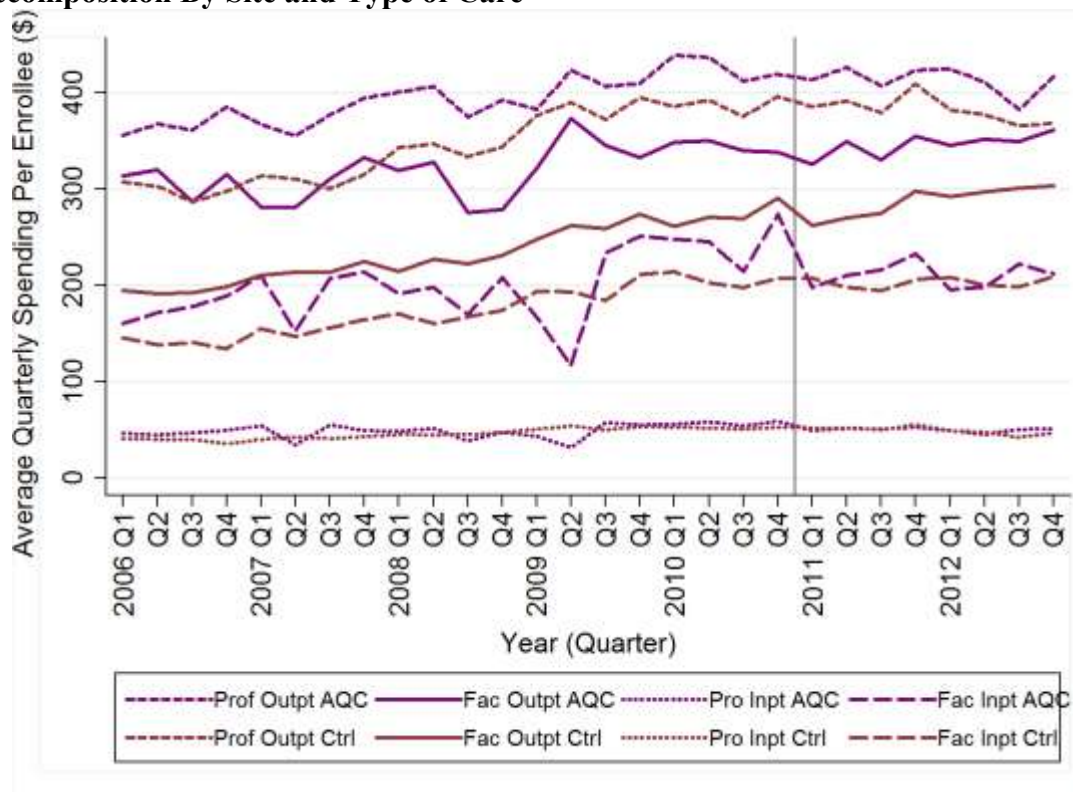
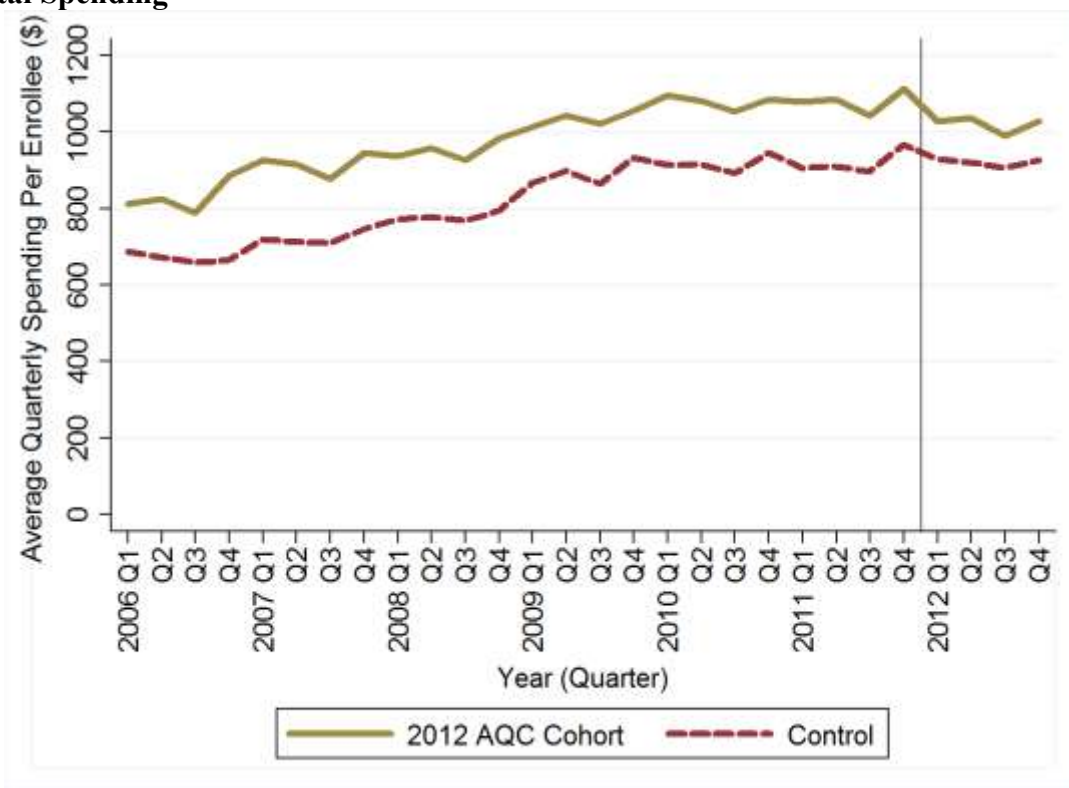


Figure S3. Unadjusted Spending: 2012 AQC Cohort vs. Control.*

A. Total Spending



B. Decomposition By Site and Type of Care

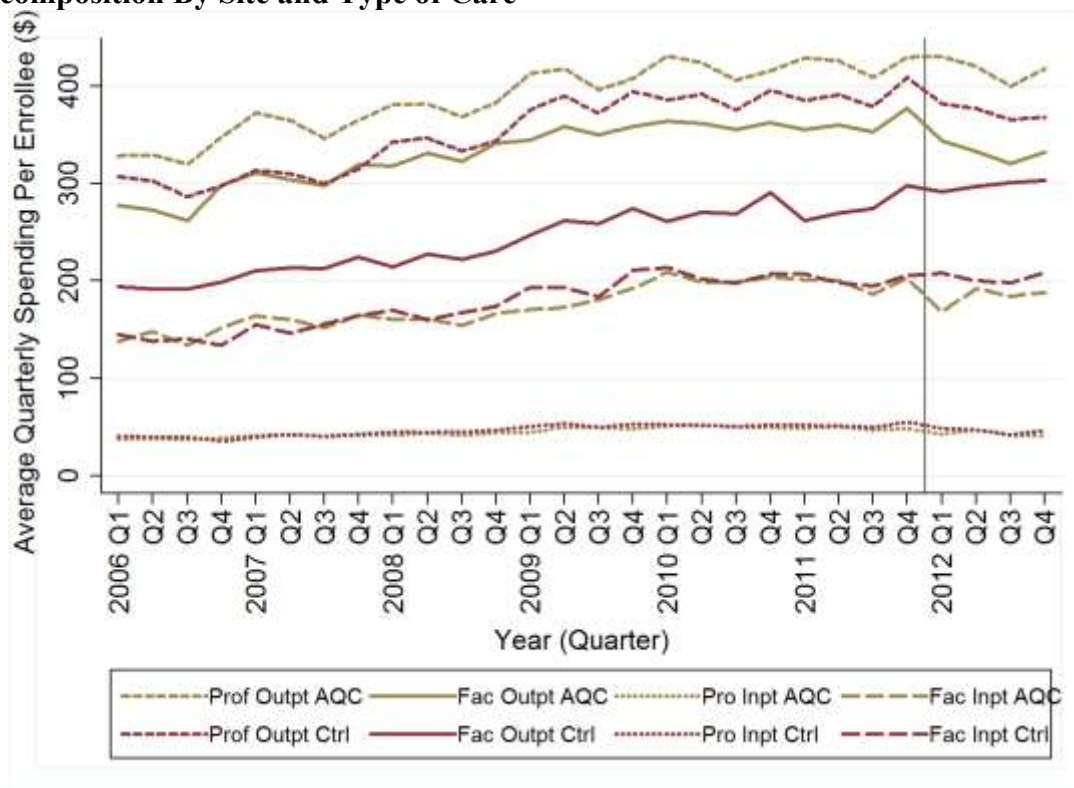
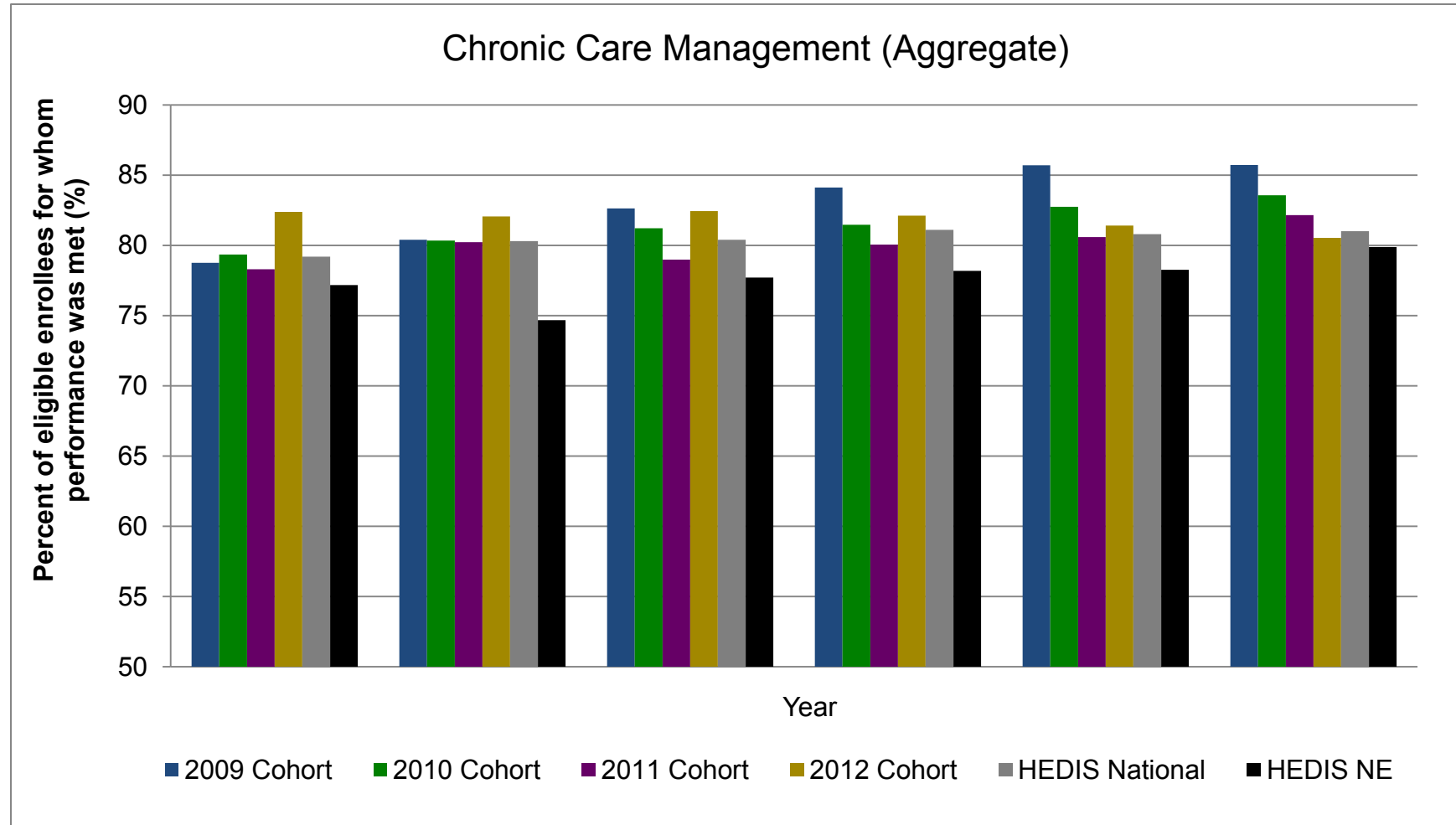
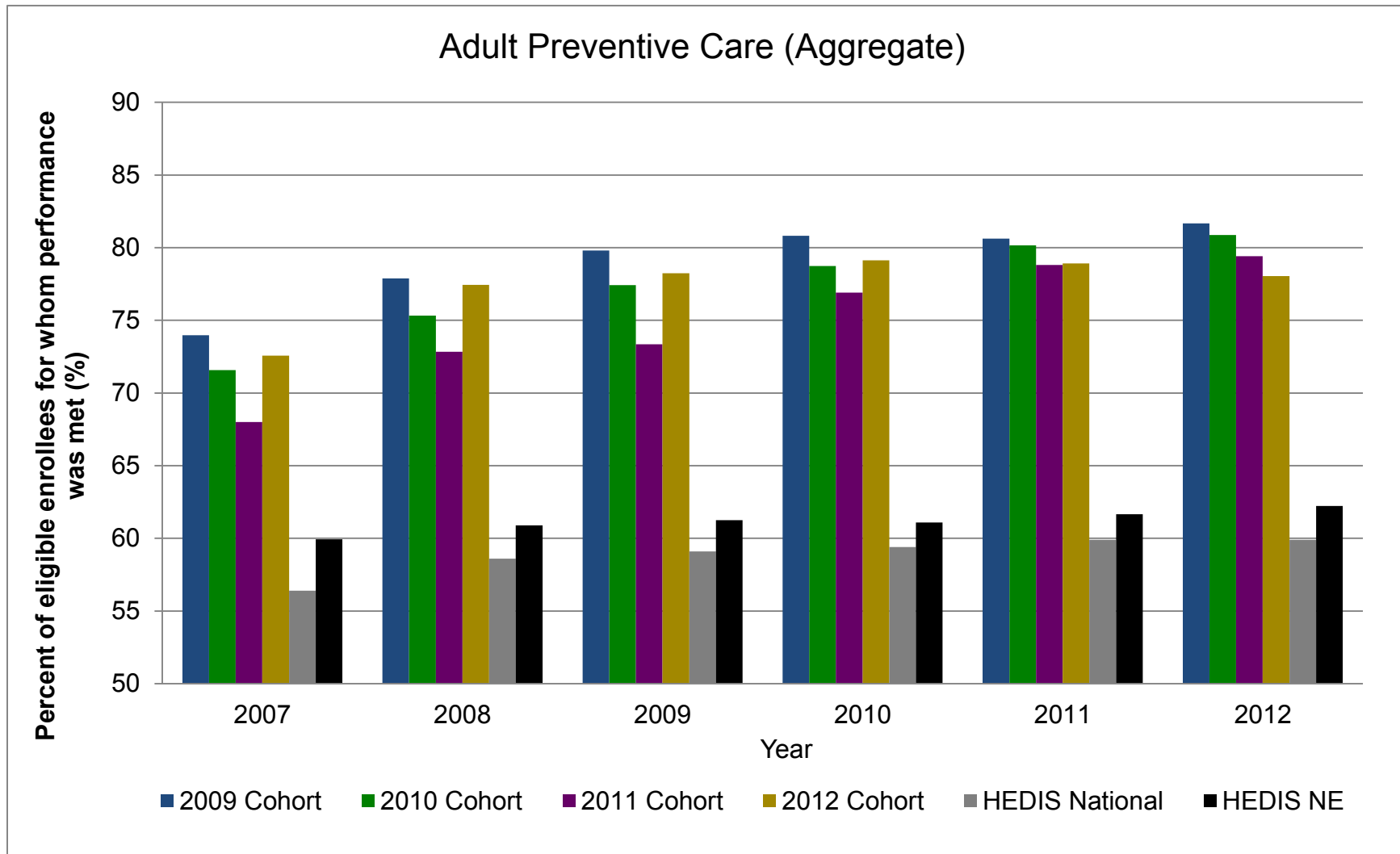


Figure S4. Process Quality by AQC Cohort, Aggregate Results 2007-2012
A. Chronic Care Management*



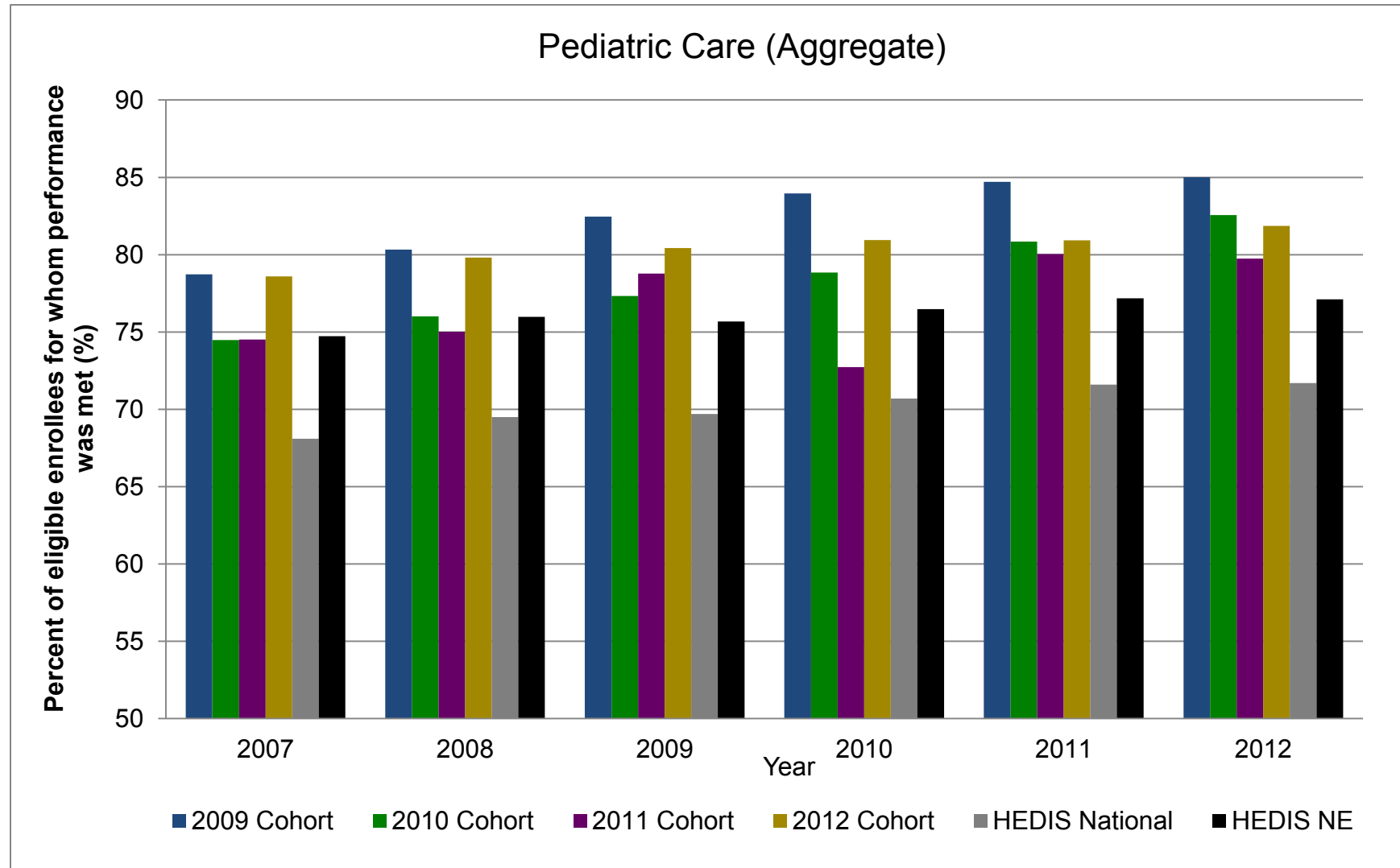
* Unadjusted performance on chronic care management quality measures for all AQC cohorts and national and New England averages in the Healthcare Effectiveness Data and Information Set (HEDIS). This aggregate measure is a weighted average of 7 individual process measures: cardiovascular low-density lipoprotein (LDL) cholesterol screening, 4 measures for enrollees with diabetes (hemoglobin A1c testing, eye exam, LDL cholesterol screening, and nephrology screening), and 2 measures for depression care (short-term prescription and maintenance prescription).

B. Adult Preventive Care*



* This aggregate measure is a weighted average of 5 individual measures: breast cancer screening, cervical cancer screening, colorectal cancer screening, chlamydia screening for enrollees 21–24 years of age, and no antibiotics for acute bronchitis.

C. Pediatric Care*



* This aggregate measure is a weighted average of 6 individual measures: Appropriate testing for pharyngitis, chlamydia screening for enrollees 16–20 years of age, no antibiotics for upper respiratory infection, and 3 measures for well child visits (well-child visits in the first 15 months of life, well-child visits in the third to sixth years of life, and adolescent well-care visits in the 12th to 21st years of life).

Table S1. Quality Measurements in the Alternative Quality Contract*

Ambulatory Measures					Hospital Measures				
	Measure	Gate 1	Gate 5 Weight			Measure	Gate 1	Gate 5	Weight
Process	Depression					AMI			
	1 Acute Phase Prescription	65.3	80.0	1.0		1 ACE/ARB for LVSD	89.1	98.9	1.0
	2 Continuation Phase Prescription	49.6	70.0	1.0		2 Aspirin at Arrival	98.3		1.0
	Diabetes					3 Aspirin at Discharge	98.2		1.0
	3 Hemoglobin A1c Testing	69.9	83.2	1.0		4 Beta Blocker at Arrival *	96.9		1.0
	4 Eye Exam	58.0	72.1	1.0		5 Beta Blocker at Discharge	98.5		1.0
	5 Nephropathy Screening	79.7	91.4	1.0		6 Smoking Cessation	93.1	99.9	1.0
	Cholesterol Management					Heart Failure			
	6 Diabetes LDL-C Screening	85.3	93.8	1.0		7 ACE LVSD	87.3	98.9	1.0
	7 Cardiovascular LDL-C Screening	85.3	93.8	1.0		8 LVS Function Evaluation	95.1	100.0	1.0
	Preventive Screening/Treatment					9 Discharge Instructions	71.4	98.5	1.0
	8 Breast Cancer Screening	77.1	90.0	1.0		10 Smoking Cessation	88.3	99.6	1.0
	9 Cervical Cancer Screening	83.5	92.4	1.0		Pneumonia			
	10 Colorectal Cancer Screening	65.2	83.3	1.0		11 Flu Vaccine	77.8	98.6	1.0
	Chlamydia Screening					13 Antibiotics within 6 Hours	95.6	99.8	1.0
	11 Ages 16-20	45.9	63.7	0.5		14 Oxygen Assessment	100.0		1.0
	12 Ages 21-24	50.1	67.3	0.5		15 Smoking Cessation	86.7	99.8	1.0
	Adult Respiratory Testing/Treatment					16 Antibiotic Selection	87.4	95.4	1.0
	13 Acute Bronchitis	Reporting Only 2009, 2010			1.0	17 Blood Culture	91.0	98.0	1.0
Medication Management					Surgical Infection				
14 Digoxin Monitoring	83.9	91.6	1.0		18 Antibiotic Received	86.5	98.9	1.0	
Pedi: Testing/Treatment					19 Received Appropriate Preventive	94.1	99.4	1.0	
15 Upper Respiratory Infection (URI)	90.6	97.7	1.0		20 Antibiotic Discontinued	77.9	96.2	1.0	
16 Pharyngitis	83.1	99.6	1.0		21 In-Hospital Mortality - Overall	2.15	0.88	1.0	
Pedi: Well Care Visits					22 Wound Infection	0.30	0.09	1.0	
17 First 15 Months of Life	91.8	99.3	1.0		23 Select Infections Due to Medical Care	0.18	0.02	1.0	
18 3-6 Years of Age	85.5	99.2	1.0		24 AMI after Major Surgery	0.55	0.10	1.0	
19 Adolescent Well Care Visits	60.0	87.7	1.0		25 Pneumonia after Major Surgery	1.57	0.60	1.0	
Outcomes	Diabetes					26 Post-Operative PE/DVT	0.93	0.22	1.0
	20 HbA1c in Poor Control	45.0	4.7	3.0		27 Birth Trauma - Injury to Neonate	0.20	0.01	1.0
	21 LDL-C Control (<100mg)	33.4	75.6	3.0		28 OB Trauma - Vaginal w/o Instrument	3.54	1.54	1.0
	22 Blood Pressure Control (130/80)	30.9	47.3	3.0		Hospital Patient Experience (H-CAHPS)			
	Hypertension					29 Communication with Nurses	72.6	81.2	1.0
	23 Controlling High Blood Pressure	71.6	82.5	3.0		30 Communication with Doctors	78.1	85.5	1.0
	Cardiovascular Disease					31 Responsiveness of Staff	58.4	76.4	1.0
24 LDL-C Control (<100mg)	33.4	75.6	3.0		32 Discharge Information	77.7	90.4	1.0	
Patient Experience	Patient Experiences - Adult								
	25 Communication Quality	91.0	98.0	1.0					
	26 Knowledge of Patients	80.0	95.0	1.0					
	27 Integration of Care	80.0	96.0	1.0					
	28 Access to Care	79.0	96.0	1.0					
	Patient Experiences - Pediatric								
	29 Communication Quality	95.0	97.0	1.0					
	30 Knowledge of Patients	89.0	93.0	1.0					
	31 Integration of Care	85.0	91.0	1.0					
	32 Access to Care	70.0	90.0	1.0					

* Performance on a quality measure is derived on an annual basis and calculated as the percent of eligible enrollees whose care met a defined threshold of quality for the measure. For example, the quality measure labeled Eye Exam reflects the percent of diabetes patients who received an eye exam in a year. Scores for all measures are weighted and summed to a total score, which is used to calculate the magnitude of bonus payments an organization receives. There are 5 pre-defined levels of overall performance (gates), and the highest level rewards the largest bonus or the most favorable ratio of shared savings to shared risk that an organization faces under the budget. Some measures are weighted differently than other measures.

Table S2. Provider Organizations in the Alternative Quality Contract*

	2009 Cohort							2010 Cohort				2011 Cohort	2012 Cohort				
Organization	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Specialists	600	20	200	250	750	100	300	750	225	0	50	1300	600	1000	150	500	5000
PCPs	400	70	50	100	350	50	100	300	100	100	50	425	300	300	75	125	1300
Hospital affiliation	Yes	Yes	Yes	Yes	No [†]	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	No	Yes	Yes	Yes

* Numbers of physicians are estimated to the nearest 25. “PCPs” denotes primary care physicians.

[†] No hospital affiliation for the first 3 years in the contract. Hospital affiliation began in 2012.

Table S3. Characteristics of the Population: AQC Cohorts vs. Control***A. 2009 AQC Cohort vs. Control**

	Pre-intervention (2006-2008)			Post-intervention (2009-2012)		
	2009 Cohort	Control Group	P Value	2009 Cohort	Control Group	P Value
Age (yr)	33.5 ± 18.2	33.4 ± 18.2	<0.001	34.6 ± 18.1	33.9 ± 18.3	<0.001
Female sex (%)	52.3	50.1	<0.001	52.1	50.0	<0.001
DxCG risk score						
Mean	0.97	0.94	<0.001	1.07	1.05	<0.001
Median	0.45	0.39		0.50	0.43	
(Interquartile range)	(0.19–1.00)	(0.14–0.97)		(0.21–1.11)	(0.16–1.09)	
Cost-sharing (%)						
Mean	10.9	17.8	<0.001	13.0	20.0	<0.001
Median	8.3	13.9		8.8	15.1	
(Interquartile range)	(4.1–14.5)	(7.9–22.9)		(4.4–16.6)	(8.5–25.8)	

B. 2010 AQC Cohort vs. Control

	Pre-intervention (2006-2009)			Post-intervention (2010-2012)		
	2010 Cohort	Control Group	P Value	2010 Cohort	Control Group	P Value
Age (yr)	34.5 ± 18.2	33.5 ± 18.3	<0.001	37.3 ± 17.5	34.0 ± 18.3	<0.001
Female sex (%)	52.1	50.2	<0.001	52.0	49.8	<0.001
DxCG risk score						
Mean	1.04	0.97	<0.001	1.15	1.06	<0.001
Median	0.49	0.40		0.55	0.43	
(Interquartile range)	(0.19–1.10)	(0.15–1.00)		(0.23–1.22)	(0.16–1.09)	
Cost-sharing (%)						
Mean	10.8	17.7	<0.001	13.2	21.0	<0.001
Median	8.1	13.8		8.8	15.8	
(Interquartile range)	(4.2–14.0)	(7.8–22.9)		(4.6–16.6)	(8.9–27.1)	

C. 2011 AQC Cohort vs. Control

	Pre-intervention (2006-2010)			Post-intervention (2011-2012)		
	2011 Cohort	Control Group	P Value	2011 Cohort	Control Group	P Value
Age (yr)	41.1 ± 14.8	33.6 ± 18.3	<0.001	41.1 ± 14.7	33.9 ± 18.2	<0.001
Female sex (%)	51.4	50.2	<0.001	53.5	49.5	<0.001
DxCG risk score						
Mean	1.24	0.98	<0.001	1.28	1.05	<0.001
Median	0.60	0.41		0.63	0.43	
(Interquartile range)	(0.24–1.32)	(0.15–1.02)		(0.26–1.38)	(0.16–1.08)	
Cost-sharing (%)						
Mean	11.9	18.1	<0.001	14.0	21.7	<0.001
Median	7.9	13.8		9.0	16.5	
(Interquartile range)	(3.9–14.8)	(7.9–23.2)		(4.5–18.0)	(9.3–28.0)	

D. 2012 AQC Cohort vs. Control

	Pre-intervention (2006-2011)			Post-intervention (2012)		
	2012 Cohort	Control Group	P Value	2012 Cohort	Control Group	P Value
Age (yr)	31.8 ± 19.1	33.6 ± 18.3	<0.001	31.9 ± 18.1	33.9 ± 18.2	<0.001
Female sex (%)	52.0	50.1	<0.001	51.4	49.5	<0.001
DxCG risk score						
Mean	1.03	1.00	<0.001	1.06	1.03	<0.001
Median	0.46	0.41		0.48	0.42	
(Interquartile range)	(0.19–1.05)	(0.16–1.03)		(0.20–1.09)	(0.16–1.06)	
Cost-sharing (%)						
Mean	10.0	18.6	<0.001	13.0	22.0	<0.001
Median	7.1	14.3		8.4	16.7	
(Interquartile range)	(3.6–12.5)	(8.1–23.9)		(4.3–16.5)	(9.3–28.3)	

* Intervention subjects were Blue Cross Blue Shield of Massachusetts enrollees who designated primary care physicians in organizations that joined the AQC. Control subjects were enrolled in private employer-sponsored plans in other Northeastern states (CT, ME, NH, NJ, NY, PA, RI, VT). The DxCG risk score is reflective enrollee health status, calculated using information from demographics and ICD-9 diagnoses. It is similar to the Hierarchical Condition Category risk score system used by Medicare. Cost sharing is the portion of spending paid by enrollees. In each comparison, p values show the significance of differences in the levels of the variables in both the pre- and post-intervention periods. Changes over time were more similar. These enrollee characteristics were adjusted for in the base statistical model and were excluded in the sensitivity analyses (Table S6 below) to test the robustness of the main results.

Table S4. Change in Average Quarterly Spending per Enrollee, AQC Cohorts vs. Control*

	AQC Enrollees in MA			Individuals in Control States			Between-Group Change		
	Pre 2006-08	Post 2009-12	Change	Pre 2006-08	Post 2009-12	Change	Unadjusted	Adjusted	P
2009 Cohort	789.35	913.15	123.8	731.61	911.4	179.79	-55.99	-62.21	<0.001
Inpt Professional	39.15	45.16	6.01	42.54	50.72	8.18	-2.17	0.40	0.82
Inpt Facility	149.05	185.58	36.53	156.33	201.1	44.76	-8.24	-3.32	0.524
Outpt Professional	340.67	384.23	43.56	319.74	383.7	63.97	-20.4	-15.35	0.004
Outpt Facility	219.91	252.65	32.74	213	275.88	62.88	-30.13	-48.67	<0.001
BETOS Category									
Eval & Mgmt	196.46	231.93	35.48	134.15	157.5	23.35	12.13	3.42	0.207
Procedures	173.78	202.45	28.66	167.83	209.04	41.21	-12.55	-17.62	<0.001
Imaging	99.14	101.03	1.89	66.24	77.01	10.77	-8.88	-10.97	<0.001
Tests	71.73	80.92	9.19	52.83	67.69	14.85	-5.66	-7.83	<0.001
DME	10.24	12.64	2.4	9.61	13.08	3.46	-1.06	-1.57	0.002
Other/Unclassified	66.86	79.09	12.23	46.24	56.17	9.93	2.3	-5.54	0.22
	Pre 2006-09	Post 2010-12	Change	Pre 2006-09	Post 2010-12	Change	Unadjusted	Adjusted	P
2010 Cohort	876.42	954.74	78.32	772.69	919.43	146.74	-68.42	-81.92	<0.001
Inpt Professional	41.02	44.5	3.48	44.96	50.24	5.28	-1.81	1.23	0.583
Inpt Facility	166.53	197.54	31.01	166.42	203.31	36.88	-5.88	6.21	0.409
Outpt Professional	354.02	383.15	29.13	336.05	383.98	47.92	-18.79	-17.86	0.007
Outpt Facility	274.18	283.81	9.63	225.25	281.9	56.65	-47.02	-80.98	<0.001
BETOS Category									
Eval & Mgmt	211.58	236.8	25.22	139.61	158.34	18.72	6.49	5.53	0.074
Procedures	193.39	219.19	25.8	177.92	209.87	31.95	-6.15	-25.42	<0.001
Imaging	107.4	103.84	-3.57	69.66	76.03	6.37	-9.94	-15.63	<0.001
Tests	82.84	82.46	-0.38	56.09	68.56	12.48	-12.85	-17.36	<0.001
DME	11.93	13.66	1.73	10.28	13.43	3.15	-1.42	-1.87	0.05
Other/Unclassified	79.53	80.13	0.6	47.94	57.47	9.52	-8.93	-13.12	0.055

	Pre 2006-10	Post 2011-12	Change	Pre 2006-10	Post 2011-12	Change	Unadjusted	Adjusted	P
2011 Cohort	1,044.91	1,070.56	25.65	797.82	920.64	122.83	-97.18	-97.10	<0.001
Inpt Professional	51.26	50.3	-0.96	46.17	49.44	3.27	-4.23	-4.49	0.225
Inpt Facility	215.15	211.39	-3.76	173.19	202.4	29.21	-32.97	-30.95	0.07
Outpt Professional	399.85	414.1	14.25	344.94	382.54	37.6	-23.35	-22.65	0.013
Outpt Facility	326.77	344.63	17.86	233.52	286.26	52.75	-34.89	-28.27	0.028
BETOS Category									
Eval & Mgmt	238.88	262.62	23.74	142.85	158.41	15.57	8.17	22.96	<0.001
Procedures	232.25	248.58	16.33	184.06	208.31	24.25	-7.93	-14.83	0.094
Imaging	135.77	124.22	-11.55	70.88	75.74	4.86	-16.41	-14.61	<0.001
Tests	101.18	101.16	-0.02	57.8	69.81	12.01	-12.03	-8.64	<0.001
DME	16.3	15.3	-1.01	10.72	13.7	2.98	-3.99	-2.04	0.074
Other/Unclassified	81.96	86.55	4.59	49.05	58.98	9.93	-5.34	-22.95	0.15
	Pre 2006-11	Post 2012	Change	Pre 2006-11	Post 2012	Change	Unadjusted	Adjusted	P
2012 Cohort	981.06	1,022.80	41.74	817.93	921.01	103.08	-61.33	-59.39	0.035
Inpt Professional	45.11	43.46	-1.65	47.16	46.32	-0.84	-0.81	5.40	0.245
Inpt Facility	173.87	183.26	9.39	177.81	203.58	25.77	-16.37	10.65	0.464
Outpt Professional	386.48	417.73	31.25	352.48	373.01	20.53	10.72	14.26	0.049
Outpt Facility	331.64	332.86	1.22	240.48	298.1	57.62	-56.4	-95.05	<0.001
BETOS Category									
Eval & Mgmt	247.44	278.84	31.4	145.34	158.83	13.48	17.91	24.23	<0.001
Procedures	213.99	230.17	16.18	188.1	207.89	19.79	-3.61	-9.98	0.392
Imaging	123	105.45	-17.55	71.39	77.75	6.36	-23.91	-22.63	<0.001
Tests	95.26	95.57	0.31	59.4	72.41	13.01	-12.71	-10.14	<0.001
DME	12.4	12.56	0.16	11.19	13.83	2.64	-2.48	-2.11	0.03
Other/Unclassified	90.51	93.04	2.53	50.26	61.9	11.64	-9.11	-23.93	0.093

Values are in units of 2012 U.S. dollars per quarter per enrollee. Adjusted estimates are derived from the statistical model as described in the Methods. “Inpt” = inpatient; “Outpt” = outpatient; “BETOS” = Berenson-Eggers Type of Service categories from CMS.

Table S5. Adjusted Spending: 2009 AQC Cohort Subgroups vs. Control*

	2009 AQC Cohort Enrollees			Individuals in Control States			Between-Group Change		
	Pre 2006-08	Post 2009-12	Change	Pre 2006-08	Post 2009-12	Change	Unadjusted	Adjusted	P
Prior Risk Subgroup	786.11	916.77	130.66	731.61	911.40	179.79	-49.13	-57.61	<0.001
Inpt Professional	39.15	45.65	6.49	42.54	50.72	8.18	-1.69	0.40	0.815
Inpt Facility	148.04	184.25	36.21	156.33	201.10	44.76	-8.56	-3.28	0.522
Outpt Professional	343.41	392.76	49.35	319.74	383.70	63.97	-14.61	-11.87	0.019
Outpt Facility	214.94	249.21	34.27	213.00	275.88	62.88	-28.61	-46.69	<0.001
BETOS Category									
Eval & Mgmt	197.53	236.98	39.45	134.15	157.50	23.35	16.10	6.13	0.018
Procedures	172.80	202.74	29.94	167.83	209.04	41.21	-11.27	-16.63	<0.001
Imaging	98.73	101.25	2.51	66.24	77.01	10.77	-8.26	-10.75	<0.001
Tests	70.93	81.07	10.14	52.83	67.69	14.85	-4.72	-7.22	<0.001
DME	10.01	12.11	2.10	9.61	13.08	3.46	-1.36	-1.86	<0.001
Other/Unclassified	66.09	78.89	12.79	46.24	56.17	9.93	2.87	-5.29	0.244
	Pre 2006-08	Post 2009-12	Change	Pre 2006-08	Post 2009-12	Change	Unadjusted	Adjusted	P
No Prior Risk Subgroup	817.97	890.51	72.53	731.61	911.40	179.79	-107.26	-68.66	<0.001
Inpt Professional	39.10	42.10	3.00	42.54	50.72	8.18	-5.19	2.44	0.26
Inpt Facility	157.92	193.90	35.98	156.33	201.10	44.76	-8.79	5.26	0.577
Outpt Professional	316.57	330.87	14.30	319.74	383.70	63.97	-49.66	-23.13	<0.001
Outpt Facility	263.76	274.21	10.45	213.00	275.88	62.88	-52.43	-64.25	<0.001
BETOS Category									
Eval & Mgmt	187.02	200.37	13.35	134.15	157.50	23.35	-10.00	-8.67	<0.001
Procedures	182.51	200.62	18.11	167.83	209.04	41.21	-23.10	-19.80	<0.001
Imaging	102.72	99.70	-3.02	66.24	77.01	10.77	-13.79	-11.00	<0.001
Tests	78.78	80.02	1.24	52.83	67.69	14.85	-13.61	-11.48	<0.001
DME	12.28	15.94	3.66	9.61	13.08	3.46	0.20	-0.38	0.776
Other/Unclassified	73.64	80.35	6.71	46.24	56.17	9.93	-3.22	-6.31	0.422

Table S6: Sensitivity Analyses**A. Alterations to the Statistical Model[†]**

	(1) Base Model	(2) Percent Cost-sharing	(3) Omit Plan Type FE	(4) Omit State FE	(5) Omit Plan FE	(6) Omit State & Plan FE	(7) Omit Age & Sex	(8) Omit Risk Score	(9) Omit Age & Risk	(10) Prior with Plan FE
AQC*Post	-62.21*** (11.12)	-63.64*** (11.96)	-62.16*** (11.12)	-62.03*** (11.14)	-66.78*** (12.53)	-62.55*** (12.96)	-65.22*** (13.43)	-63.36*** (16.82)	-49.16*** (18.06)	-59.13*** (15.45)
AQC	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Years	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y		Y		
Sex	Y	Y	Y	Y	Y	Y		Y		
Risk	Y	Y	Y	Y	Y	Y	Y			
State FE	Y	Y	Y		Y					
Plan type	Y	Y								
Plan FE	Y		Y	Y						Y
% CS		Y								
Observations	3,715,260	3,715,048	3,715,260	3,715,260	3,715,260	3,715,260	3,715,260	3,729,885	3,729,885	3,729,885
R-squared	0.529	0.528	0.529	0.529	0.528	0.528	0.527	0.015	0.001	0.005

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[†] These sensitivity analyses test the robustness of our main results against various changes in the model. Column 1 reproduces the main coefficient of interest (average quarterly change in spending associated with the AQC over the first 4 years of the contract, using the 2009 cohort vs. control comparison). The remaining columns show the same coefficient in alternative scenarios: (2) percent cost sharing in place of plan fixed effects; (3) exclusion of plan type fixed effects; (4-5) exclusion of state or plan fixed effects; (6) exclusion of state and plan fixed effects; (7) exclusion of age and sex; (8) exclusion of risk score; (9) exclusion of age, sex, and risk score; (10) exclusion of age, sex, and risk score with inclusion of plan fixed effects. CS is cost sharing; it is derived by calculating the percent of spending paid by the enrollee out of pocket for the 10 most frequent services and then averaging those percentages by plan. This is a reflection of plan generosity. State FE are state fixed effects. Plan FE are plan fixed effects, where the plan is a unique plan number or benefit design issued by a given insurer, rather than a unique insurer. The statistical model is described in the text of the paper.

B. Alterations in Variables or Sample^{††}

	Tests of alternative, non-preferred control groups										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Base Model	Risk Score Deciles	Net of Cost Sharing	With Rx Drugs	Risk Prosp Lagged	Seven-Year Continu	Enrollee-Quarter Model	HMO Controls Only	MA Only Controls	Non-AQC BCBSMA Controls	Non-MA USA Controls
AQC*Post	-62.21*** (11.12)	-91.12*** (11.89)	-56.65*** (10.84)	-82.12*** (14.15)	-48.46** (18.52)	-84.14** (17.24)	-54.51*** (13.11)	-47.65*** (11.14)	-32.47** (13.41)	-17.82*** (6.19)	-66.29** (28.66)
AQC	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Years	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sex	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Risk	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Plan type	Y	Y	Y	Y	Y	Y	Y		Y	Y	Y
Plan FE	Y	Y	Y	Y	Y	Y	Y	Y		Y	Y
Observations	3,715,260	3,729,885	3,677,771	3,715,260-	2,483,522	599,247	14,053,576	2,330,377	2,218,738	3,514,405	7,838,725
R-squared	0.529	0.160	0.523	0.543	0.145	0.499	0.240	0.545	0.527	0.536	0.015

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

^{††} These sensitivity analyses test the robustness of our main results against changes in the variables or sample. Column 1 is again the main coefficient of interest. The remaining columns show the following modifications: (2) risk scores in deciles rather than a continuous variable; (3) excluding cost sharing from spending; (4) including prescription drug spending; (5) prospective risk score lagged by 1 year; (6) restricting to continuous enrollees over 7 years during the study period; (7) quarterly model at the enrollee level. Columns (8-11) test alternative control groups that were possible to construct using the available data. These alternative control groups have drawbacks that we describe here and note in the paper. (8) HMO only controls from the 8 Northeastern states. This group fails to capture all enrollees in plans comparable to the AQC, which require designating a PCP and have incentives for receiving care in network. Also, this group had significant differences in pre-intervention spending trends compared to the AQC. (9) All Massachusetts control group. This group is not ideal because it contains BCBSMA (treatment) enrollees as well; we could not separate BCBSMA enrollees from Harvard Pilgrim, Tufts, or other private payers in MA due to the absence of payer IDs in the Truven data for confidentiality. Moreover, this control group also had significant differences in pre-intervention spending trends relative to the AQC. (10) Non-AQC BCBSMA control group (enrollees whose providers had not joined the AQC by 2012). This is not an ideal control group because the remaining providers in non-incentive contracts were small, rural practices that received lower fee updates from BCBSMA as a consequence of remaining in fee-for-service. Moreover, this control group also had significant differences in pre-intervention spending trends relative to the AQC. The Massachusetts only control groups are also susceptible to spillover effects. (11) National controls comprising a 10% random sample of enrollees in the 49 non-Massachusetts states in the Truven data. As with the main control group, national controls are susceptible to other factors in Massachusetts affecting the results, which we discuss in the paper. However, the national control group does not contaminate controls with treatment subjects and is less susceptible to AQC spillover effects within Massachusetts. Of note, similar to the baseline control group, this national control group demonstrated no significant differences in pre-intervention spending trends relative to the AQC.

Table S7. HEDIS New England Averages***A. Process Measures***Chronic Care Management*

	HEDIS New England Averages					
	2007	2008	2009	2010	2011	2012
Cardiovascular LDL cholesterol screening	90.4	91.5	91.6	91.7	90.5	90.2
Diabetes - HbA1c testing	91.9	92.3	92.1	92.3	92.0	92.7
Diabetes - Eye exams	70.8	71.7	70.3	70.5	69.2	70.2
Diabetes - LDL screening	87.0	87.9	87.5	88.0	87.3	87.4
Diabetes - Nephrology screening	83.4	85.5	85.4	85.9	85.3	87.1
Depression - Short term prescription	66.1	42.6	66.6	67.3	69.9	73.2
Depression - Maintenance prescription	50.6	51.2	50.4	51.6	53.6	58.4
Average	77.2	74.7	77.7	78.2	78.3	79.9

Adult Preventive Care

	HEDIS New England Averages					
	2007	2008	2009	2010	2011	2012
Breast cancer screening	76.7	77.2	78.4	77.5	76.8	77.8
Cervical cancer screening	86.1	84.9	81.9	81.2	80.4	79.2
Colorectal cancer screening	67.6	69.4	71.5	71.9	70.7	71.5
Chlamydia screening for 21-24 year-olds*	48.1	51.2	53.6	54.7	57.1	56.9
No antibiotics for acute bronchitis	21.2	21.8	20.8	20.1	23.3	25.7
Average	59.9	60.9	61.2	61.1	61.7	62.2

* Specification changed to 21-24 year-olds in 2008 and beyond; age range was 21-25 in 2007.

Pediatric Care

	HEDIS New England Averages					
	2007	2008	2009	2010	2011	2012
Appropriate testing for pharyngitis	81.4	83.8	85.4	87.8	89.0	89.4
Chlamydia screening 16-20 years of age	45.4	47.9	47.4	47.7	49.0	47.2
No antibiotics for upper respiratory infection	91.1	91.8	92.1	92.2	92.1	91.6
Well child visits first 15 months of life	85.5	86.4	83.4	85.3	85.1	86.3
Well child visits third to sixth years of life	84.1	84.4	84.9	85.0	85.8	86.1
Adolescent well care visits 12th to 21st years of life	60.9	61.6	60.9	60.9	62.1	62.1
Average	74.7	76.0	75.7	76.5	77.2	77.1

B. Outcome Measures

	HEDIS					
	New England Average					
	2007	2008	2009	2010	2011	2012
Diabetic HbA1c Control (≤ 9 percent)	77.5	77.7	76.6	77.4	74.9	75.8
Diabetic LDL Cholesterol Control (<100mg/dl)	49.1	49.5	49.8	50.8	51.4	52.3
Diabetic Blood Pressure Control (<140/80 mmHg) [†]	35.7	37.1	36.4	Not reported	47.6	50.5
LDL Cholesterol Control in Patients with Coronary Artery Disease (<100mg/dl)	66.0	66.6	65.8	65.3	64.7	65.0
Blood Pressure Control in Patients with Hypertension (<140/90 mmHg)	66.7	68.0	68.5	68.1	68.2	68.2
Average	59.0	59.8	59.4	--	61.4	62.4

* Values designate the percent of eligible enrollees for a measure whose care achieved a defined threshold of quality performance for the measure. “HEDIS” is the Healthcare Effectiveness Data and Information Set. “HbA1c” is hemoglobin A1c. “LDL” is low-density lipoprotein cholesterol.

[†] Specification for this measure was <130/80 from 2007-2009, changed to <140/80 in 2010 and beyond. Figure 2 in the main paper plots the averages seen here. For 2010, in which the diabetic blood pressure measure was not reported, a value for that measure was imputed using the linear average of 2009 and 2011 values, which was then averaged with the rest of the 2010 measures to produce an overall 2010 average.

Table S8. Outcome Quality: AQC Cohorts vs. HEDIS National Average**A. 2009 AQC Cohort vs. Control**

	2009 AQC Cohort				HEDIS National Average	
	2009	2010	2011	2012	2011	2012
Outcome Measures	<i>Percent of population achieving performance (%)</i>				<i>Percent of population achieving performance (%)</i>	
Diabetic HbA1c Control (≤ 9 percent)	80.6	82.4	83.9	84.1	71.7	71.5
Diabetic LDL Cholesterol Control (< 100 mg/dl)	57.7	61.0	64.1	65.2	48.1	48.4
Diabetic Blood Pressure Control ($< 140/80$ mmHg)	51.6	54.8	60.6	65.4	44.2	44.3
LDL Cholesterol Control in Patients with Coronary Artery Disease (< 100 mg/dl)	69.9	72.3	74.0	74.8	59.8	59.9
Blood Pressure Control in Patients with Hypertension ($< 140/90$ mmHg)	68.4	71.1	78.3	80.4	65.4	63.0
Average	65.6	68.3	72.2	74.0	57.8	57.4

B. 2010 AQC Cohort vs. Control

	2010 AQC Cohort				HEDIS National Average	
	2009	2010	2011	2012	2011	2012
Outcome Measures	<i>Percent of population achieving performance (%)</i>				<i>Percent of population achieving performance (%)</i>	
Diabetic HbA1c Control (≤ 9 percent)	N/A	80.0	81.9	84.7	71.7	71.5
Diabetic LDL Cholesterol Control (< 100 mg/dl)	N/A	55.8	57.6	61.0	48.1	48.4
Diabetic Blood Pressure Control ($< 140/80$ mmHg)	N/A	54.1	59.1	63.3	44.2	44.3
LDL Cholesterol Control in Patients with Coronary Artery Disease (< 100 mg/dl)	N/A	66.0	67.7	71.2	59.8	59.9
Blood Pressure Control in Patients with Hypertension ($< 140/90$ mmHg)	N/A	74.6	78.5	81.9	65.4	63.0
Average	N/A	66.1	69.0	72.4	57.8	57.4

C. 2011 AQC Cohort vs. Control

	2011 AQC Cohort				HEDIS National Average	
	2009	2010	2011	2012	2011	2012
Outcome Measures	<i>Percent of population achieving performance (%)</i>				<i>Percent of population achieving performance (%)</i>	
Diabetic HbA1c Control (≤ 9 percent)	N/A	80.2	78.8	80.5	71.7	71.5
Diabetic LDL Cholesterol Control (< 100 mg/dl)	N/A	54.3	54.3	56.8	48.1	48.4
Diabetic Blood Pressure Control ($< 140/80$ mmHg)	N/A	38.3	51.3	52.6	44.2	44.3
LDL Cholesterol Control in Patients with Coronary Artery Disease (< 100 mg/dl)	N/A	68.2	67.1	73.2	59.8	59.9
Blood Pressure Control in Patients with Hypertension ($< 140/90$ mmHg)	N/A	69.9	67.0	70.9	65.4	63.0
Average	N/A	62.2	63.7	66.8	57.8	57.4

D. 2012 AQC Cohort vs. Control

	2012 AQC Cohort				HEDIS National Average	
	2009	2010	2011	2012	2011	2012
Outcome Measures	<i>Percent of population achieving performance (%)</i>				<i>Percent of population achieving performance (%)</i>	
Diabetic HbA1c Control (≤ 9 percent)	N/A	82.4	83.9	84.1	71.7	71.5
Diabetic LDL Cholesterol Control (< 100 mg/dl)	N/A	61.0	64.1	65.2	48.1	48.4
Diabetic Blood Pressure Control ($< 140/80$ mmHg)	N/A	54.8	60.6	65.4	44.2	44.3
LDL Cholesterol Control in Patients with Coronary Artery Disease (< 100 mg/dl)	N/A	72.3	74.0	74.8	59.8	59.9
Blood Pressure Control in Patients with Hypertension ($< 140/90$ mmHg)	N/A	71.1	78.3	80.4	65.4	63.0
Average	N/A	68.3	72.2	74.0	57.8	57.4

* Values designate the percent of eligible enrollees for a measure whose care achieved a defined threshold of quality performance for the measure. “HEDIS” is the Healthcare Effectiveness Data and Information Set. “HbA1c” is hemoglobin A1c. “LDL” is low-density lipoprotein cholesterol. Outcomes data were not systematically collected for the 2010-2012 AQC cohorts prior to 2010.